



RESEARCH PROGRAM ON
Climate Change,
Agriculture and
Food Security



Data Ownership Agreement- Template

January 2018



Stats4SD

This document was produced in collaboration with [Statistics for Sustainable Development](#) and is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#).

Introduction

[This template assumes that:

- *A contract has been signed between partner institutions;*
- *Any intellectual property rights have been established in the contracts signed between institutions;*
- *Individual scientists and other data generators are subject to the terms of the contract signed by their institution for this project;*
- *The project is publicly funded or that a decision for making the data accessible to the public has been made by the leading institution.]*

This is an agreement between the signing institutions and scientists to establish data ownership and the right to use of data generated by the research activities of the research project *<project name>*.

It is understood that scientists and technicians working for the signing institutions are subjected to this agreement.

The agreement is based on the following principles:

- Research datasets belong to institutions not individuals as it is only institutions that can ensure long term security for data;
- Data generated by collaborating institutions belong jointly to those institutions;
- Data collected using public funds are public property; everyone has a responsibility to ensure the maximum value is realised from them;
- Scientists generating research data have a right to recognition for their work;
- Scientists generating research data using public funds have a duty to use the data for the purposes for which funding was provided and to publish the findings;
- Data ownership and intellectual property should be managed in a way to balance the interests of individual scientists, their institutions, the donors and society as a whole.

In practice these principles will be implemented as follows:

1. Institution *<institution name>* is responsible for co-ordination of all data management, sharing and accessibility activities. It will maintain databases of all data collected by the project and make the information accessible, according to the points below.
2. Each partner in this research project/programme agrees to supply all primary data and associated metadata collected as part of the project, to the co-ordinating institution for storage and archiving according to agreed activity milestones.

[Activity milestones would vary according to the activity: for a short-term single activity it might be after data entry; for longer-term activities it might be after the data entry for each site.]

The co-ordinating institution will provide instructions and support on the practical aspects of this process.

3. Primary data will be made publicly available by the coordinating institution <specify time> months after its generation.

[For single data collection activities, such as a household survey, the data could be made available 12 months post generation. For long-term studies where data collection may be longitudinal, discussions should take place between the partners to decide whether data should be made publicly available at regular intervals throughout the collection process, i.e. every 12 months, or whether the data are not made public until the end of the data collection process, which may be several years. This is likely to be established by agreements established in the main contract between institutions.]

4. Before the primary data are made public, project scientists will have a reasonable period of time in which to use the data for the agreed purposes.
5. The co-ordinating institution will safeguard the data and ensure that the scientists that have the right to access the data can do so both before and after the public release date.
6. Primary data will be made publicly available, ensuring that the generator/s of the data (the individual/s and associated institutions) is/are acknowledged as the author/s of the data. The data and the information about authorship of the data should be kept together.

Definitions

Agreed purposes

At the start of any project activity a protocol should be produced; this document includes details about the data that are to be collected and how they will be used (the agreed purposes), along with timelines for the activity which should include data collection timelines. All project partners should review and agree on the protocol prior to starting the activity.

Data

‘Data’ in this template refers not only to: datasets, metadata, observational data, and statistical data, but also items such as: survey questionnaires, reports, videos, images, maps and audio recordings.

Metadata

Metadata is information that fully describes the source and content of each dataset; it is information that allows a user to fully understand the dataset or resource. It also includes project level information that helps researchers to locate the data in a public archive. It is often described as “data about data”. For further information please see the document “Introduction to Metadata”.

Primary data

By ‘Primary data’ we mean data that are of satisfactory quality to be of use. For numerical data this means the data should be clean and may include derived variables; for textual data such as reports, they should have been through the review process by the partners involved in the project. Primary data should also be anonymised as appropriate. It is the primary data rather than the raw data that will be archived. For further information see the document “Transition from Raw to Primary Data”.